

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): HOSOYA, Mutsumi  
Serial No.: Not yet assigned  
Filed: February 2, 2004  
Title: DATA TRANSFER METHOD AND DISK CONTROL UNIT  
USING IT  
Group: Not yet assigned

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

February 2, 2004

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2003-353219, filed October 14, 2003.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



---

Carl I. Brundidge  
Registration No. 29,621

CIB/alb  
Attachment  
(703) 312-6600

日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application: 2003年10月14日

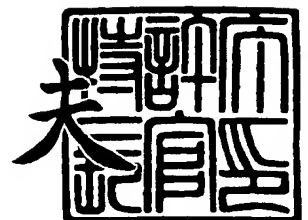
出願番号  
Application Number: 特願2003-353219  
[ST. 10/C]: [JP2003-353219]

出願人  
Applicant(s): 株式会社日立製作所

2003年10月29日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



出証番号 出証特2003-3089627

【書類名】 特許願  
【整理番号】 NT03P0585  
【提出日】 平成15年10月14日  
【あて先】 特許庁長官 殿  
【国際特許分類】 G06F 13/00  
【発明者】  
    【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所  
                                システム開発研究所内  
    【氏名】 細谷 睦  
【特許出願人】  
    【識別番号】 000005108  
    【氏名又は名称】 株式会社日立製作所  
【代理人】  
    【識別番号】 100068504  
    【弁理士】  
    【氏名又は名称】 小川 勝男  
    【電話番号】 03-3661-0071  
【選任した代理人】  
    【識別番号】 100086656  
    【弁理士】  
    【氏名又は名称】 田中 恭助  
    【電話番号】 03-3661-0071  
【選任した代理人】  
    【識別番号】 100094352  
    【弁理士】  
    【氏名又は名称】 佐々木 孝  
    【電話番号】 03-3661-0071  
【手数料の表示】  
    【予納台帳番号】 081423  
    【納付金額】 21,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1

**【書類名】 特許請求の範囲****【請求項 1】**

イニシエータからターゲットへのデータ転送の際、前記ターゲットで受信されたデータの通信エラーの有無を前記データに付加されているエラーチェックコードを用いて確認し、前記通信エラーの有無を前記ターゲットから前記イニシエータに転送ステータスとして返送し、前記転送ステータスにより前記データ転送の際に転送エラーの発生が判明した場合、前記イニシエータから前記ターゲットに前記データの再送を行うリライアブル転送において、

前記イニシエータと前記ターゲット間の前記データ転送単位である論理レコードのデータ転送方法であって、

前記論理レコードを複数まとめて一括転送し、

前記イニシエータでは、前記一括転送の単位で前記転送ステータスの確認を行い、

前記イニシエータが発行した転送リクエストにより前記論理レコードの転送がなされ、

前記ターゲットでは、前記論理レコードの前記転送リクエストに対応した完了ステータスを予め定められた一括転送条件に合致した前記論理レコードについて、その正常受信が完了した時点で前記ターゲット内に存在する完了キューに通知することを特徴とするデータ転送方法。

**【請求項 2】**

イニシエータからターゲットへのデータ転送の際、前記ターゲットで受信されたデータの通信エラーの有無を前記データに付加されているエラーチェックコードを用いて確認し、前記通信エラーの有無を前記ターゲットから前記イニシエータに転送ステータスとして返送し、前記転送ステータスにより前記データ転送の際に転送エラーの発生が判明した場合、前記イニシエータから前記ターゲットに前記データの再送を行うリライアブル転送において、

前記イニシエータと前記ターゲット間の前記データ転送の単位である論理レコードのデータ転送方法であって、

前記論理レコードを複数まとめて一括転送し、

前記イニシエータが発行した転送リクエストにより前記論理レコードの転送がなされ、それが正常に前記ターゲットに到着した時点で、前記論理レコードの前記転送リクエストに対応した完了ステータスを前記ターゲット内に存在する完了キューに通知するようになし、

前記イニシエータでは、前記一括転送の単位で前記転送ステータスの確認を行い、

前記ターゲットでは、前記一括転送の途中で前記論理レコードの転送エラーを前記エラーチェックコードにより検出した際、前記一括転送の終了まで、当該論理レコードおよびそれ以降の論理レコードの受信を無視し、前記完了ステータスの通知を抑制することを特徴とするデータ転送方法。

**【請求項 3】**

イニシエータからターゲットへのデータ転送の際、前記ターゲットで受信されたデータの通信エラーの有無を前記データに付加されているエラーチェックコードを用いて確認し、前記通信エラーの有無を前記ターゲットから前記イニシエータに転送ステータスとして返送し、前記転送ステータスにより前記データ転送の際に転送エラーの発生が判明した場合、前記イニシエータから前記ターゲットに前記データの再送を行うリライアブル転送において、

前記イニシエータと前記ターゲット間の前記データ転送の単位である論理レコードのデータ転送方法であって、

前記論理レコードを複数まとめて一括転送し、

前記イニシエータでは、前記一括転送の単位で前記転送ステータスの確認を行い、

前記イニシエータが発行した転送リクエストにより前記論理レコードの転送が正常に前記ターゲットに到着した時点で、前記論理レコードの前記転送リクエストに対応した完了ステータスを前記ターゲット内に存在する完了キューに通知し、

前記ターゲットでは、前記一括転送の途中で前記論理レコードの転送エラーを前記エラーチェックコードにより検出した際、前記一括転送の終了まで、当該論理レコード及び、それ以降の論理レコードのうち一括転送条件フィールドで指定されている論理レコードの受信を無視し、前記完了ステータスの通知を抑制することを特徴とするデータ転送方法。

【請求項 4】

前記イニシエータと前記ターゲットとの間の前記一括転送において、  
前記ターゲットでは転送エラーが検出された最初の論理レコードの ID を、前記一括転送の単位ごとに確認が行われる前記転送ステータスの中に含め、  
前記イニシエータでは前記転送ステータスをもとに、転送エラーの発生した論理レコードから再送を開始することを特徴とする請求項 1 乃至請求項 3 のいずれかに記載のデータ転送方法。

【請求項 5】

前記イニシエータと前記ターゲットとの間の前記一括転送において、  
前記ターゲットでは転送エラーが検出された論理レコードの ID のリストを、前記一括転送の単位ごとに確認が行われる前記転送ステータスの中に含め、  
前記イニシエータでは前記リストをもとに、転送エラーが発生した論理レコードの再送を行うことを特徴とする請求項 1 乃至請求項 3 のいずれかに記載のデータ転送方法。

【請求項 6】

前記一括転送の途中で、前記イニシエータ、もしくは、前記ターゲットからキャンセルリクエストを発行して、前記一括転送を中止することが可能であることを特徴とする請求項 1 乃至請求項 5 のいずれかに記載のデータ転送方法。

【請求項 7】

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部とを有し、前記ホストインターフェース部は、前記ホストコンピュータに対しリード／ライトされるデータを一時的に格納するメモリを有し、前記メモリと前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリを有し、前記メモリと前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のホストインターフェース部と前記複数のディスクインターフェース部との間の転送に請求項 1 乃至請求項 6 のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

【請求項 8】

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部とを有し、前記ホストインターフェース部は、前記ホストコンピュータに対しリード／ライトされるデータを一時的に格納するメモリを有し、前記メモリと前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリを有し、前記メモリと前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のホストインターフェース部間の転送に請求項 1 乃至請求項 6 のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

【請求項 9】

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部とを有し、前記ホストインターフェース部は、前記ホストコンピュータに対しリード／ライトされるデータを一時的に格納するメモリを有し、前記メモリと前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリを有し、前記メモリと前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のディスクインターフェース部間の転送に請求項1乃至請求項6のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

**【請求項 10】**

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、複数のメモリ部と、複数のプロセッサ部とを有し、前記ホストインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のホストインターフェース部あるいは前記複数のディスクインターフェース部と前記メモリ部との間の転送に請求項 1 乃至請求項 6 のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

**【請求項 1 1】**

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、複数のメモリ部と、複数のプロセッサ部とを有し、前記ホストインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のホストインターフェース部あるいは前記複数のディスクインターフェース部と前記プロセッサ部間の転送に請求項 1 乃至請求項 6 のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

**【請求項 1 2】**

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、複数のメモリ部と、複数のプロセッサ部とを有し、前記ホストインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のメモリ部とプロセッサ部間の転送に請求項1乃至請求項6のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

**【請求項 1 3】**

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、複数のメモリ部と、複数のプロセッサ部とを有し、前記ホストインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のメモリ部間の転送に請求項1乃至請求項6のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

【請求項 14】

ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、複数のメモリ部と、複数のプロセッサ部とを有し、前記ホストインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、

前記複数のプロセッサ部間の転送に請求項 1 乃至請求項 6 のいずれかに記載のデータ転送方法が適用されていることを特徴とするディスク制御装置。

**【書類名】 明細書****【発明の名称】** データ転送方法、および、これを用いたディスク制御装置**【技術分野】****【0001】**

本発明は、ネットワークを介して高信頼かつ高速にデータを転送するためのデータ転送方法と複数の磁気ディスク装置を制御するためのディスク制御装置に関する。

**【背景技術】****【0002】**

情報通信インフラの高度化に伴い、情報通信システムに求められる処理性能は益々高まってきている。とくに最近では、集積LSIの微細化が進展した結果、LSI間のデータ転送性能が、システム性能を左右する状況となっている。このため、IOシステムの高性能化・多機能化に対する研究が精力的に進められており、転送速度を向上するとともに、多様なトランスポート機能を持った通信プロトコル・エンジンが開発されるようになってきた。

**【0003】**

例えば、InfiniBand Architecture Specification Release 1.0aで規定されているInfiniBand転送方式は、送信キュー（SQ）および受信キュー（RQ）で構成されるキューペア（QP）と、キューペアに積まれたリクエスト処理完了時点で完了ステータスの通知される完了キュー（CQ）とで、アプリケーションプロセスとIOシステムとのインターフェースがとられるデータ転送方法となっている。その様子を図4で説明する。

**【0004】**

プロセス51とプロセス52は、それぞれ2つのキューペアを用いて通信している。プロセス51は、送信キュー11と受信キュー21からなるキューペア41と送信キュー12と受信キュー22からなるキューペア42を持っている。同様に、プロセス52は、送信キュー13と受信キュー23からなるキューペア43と送信キュー14と受信キュー24からなるキューペア44を持っている。完了キュー31には、キューペア41およびキューペア42の完了ステータスが格納され、完了キュー32には、キューペア43およびキューペア44の完了ステータスが格納される。

**【0005】**

送信キューの各エントリには、転送リクエストがおかれる。この転送リクエストにより転送されるデータの単位を論理レコードと呼ぶ。受信キューの各エントリには、受信バッファへのポインタが格納される。送信キュー12の転送リクエストは、プロセスバッファ71内のレコードバッファ81へのポインタを有し、送信キュー14の転送リクエストは、プロセスバッファ72内のレコードバッファ82へのポインタを有している。同様に、受信キュー22、24は、レコードバッファ81、82へのポインタが格納されている。

**【0006】**

通信を行う2つのキューペア間では、送信キューと受信キュー間で接続が張られる。送信キュー12は受信キュー24と、送信キュー14は受信キュー22と接続がなされている。このとき、送信キュー12に積まれた転送リクエストが処理されると、レコードバッファ81に格納されていた論理レコードが受信キュー24で指定されているレコードバッファ82へ転送される。正常な転送が完了した時点で、受信キュー24からは完了キュー32に完了ステータスが通知され、送信キュー12からは完了キュー31へ完了ステータスが通知される。

**【0007】**

これらのキューペアや完了キューの処理は、ホストチャネルアダプタ（HCA）とよばれるハードウェアで実行される。その構成例を図5に示した。HCAは、受信ポート613と送信ポート623、受信側のリンク層論理631、トランスポート層論理632、プロセッサ633、送信側のリンク層論理641、トランスポート層論理642、プロセッサ643、メモリ650、および、接続インターフェース660を有している。HCAとアプリケーションプロセスとの通信は、接続インターフェース、および、メモリを介して

行われる。受信側と送信側は、並行して動作することができ、各プロセッサと、リンク層・トランスポート層論理で、高機能プロトコル処理を高速に実行することができるようにになっている。

#### 【0008】

2つのHCA間で単純転送リクエストが処理される様子を図6で説明する。HCA1側のプロセスバッファ73内のレコードバッファ83をHCA2側のプロセスバッファ74内のレコードバッファ84に転送する場合を考える。HCA1では、レコードバッファ83を送信に適したサイズに分解し、それぞれ適当なヘッダとエラーチェックコード(CRC)を付加してパケット401-403として転送する。HCA2では、受信パケットにエラーがないかCRCで確認し、エラーを検出した場合、HCA1にその旨NAK (Negative Acknowledgement)で通知する。HCA1では、NAKが返されたパケットについては再送を行う。HCA2では、すべてのパケットが正常に受け取れた時点で、論理レコードの再構築を行い、レコードバッファ84に格納して、完了キュー34に完了ステータスを通知するとともに、HCA1に受信完了を知らせる。HCA1では、HCA2の受信完了通知を受けて、完了キュー33に転送完了ステータスを通知することで、一連の転送リクエスト処理が完了する。

#### 【0009】

別の例として、2つのHCA間でのRDMA転送リクエストが処理される様子を図7で説明する。RDMA転送では、イニシエータのアプリケーションメモリ空間75内のRDMA転送空間85をターゲットのアプリケーションメモリ空間76内のRDMA転送空間86に転送する。RDMA転送では、転送先のアプリケーションメモリ空間内に直接転送を行うので、転送先のメモリアドレス情報を付加する必要がある。それ以外は、単純転送リクエストと同様の動作をする。HCA1ではRDMA転送空間85を適当なサイズでパケットに構成し、順次HCA2に転送する。HCA2では、受信パケットをRDMA転送空間86内の決められた場所に格納し、必要ならパケットの再送処理をして、空間全体の再構築を行う。HCA2では、すべてのパケットを正常に受け取った時点で、完了キュー36に完了ステータスを通知するとともに、HCA1に受信完了を知らせる。HCA1では、HCA2の受信完了通知を受けて、完了キュー35に転送完了ステータスを通知することで、一連の転送リクエスト処理が完了する。

#### 【0010】

ここまで説明したデータ転送方法は、転送エラーの無いことを保証するリライアブルデータ転送方法であり、広く一般的に使われている基本的なものである。従来のリライアブルデータ転送方法には、以下の2つの特徴がある。

1. ターゲットでは、転送リクエスト単位である論理レコード全体にエラーの無いことを確認してから完了ステータスを通知する。
2. イニシエータでは、ターゲットからの論理レコード全体の正常転送完了ステータス通知を確認してから次の論理レコードの転送を開始する。

#### 【0011】

上記特徴について、図2、および、図3を用いて説明する。図2では、HCA1側のアプリケーション(AP)1から転送リクエスト121によって論理レコード221の転送を開始している。HCA2で転送エラーを検出した場合には再送を行う。HCA2は、正常に論理レコードの受信を完了した時点で、アプリケーション2の完了キューに完了ステータス321を通知する。完了ステータス321を受けて、アプリケーション2は、論理レコード221を使った処理721を開始することができる。論理レコード221の正常受信を完了したHCA2は、受信完了をHCA1に知らせ、HCA1はアプリケーション1の完了キューに転送完了ステータス361を通知する。この例で明らかなように、ターゲット側のアプリケーション2は、論理レコード221全体の受信が完了してから完了ステータス321を受け取る。また、イニシエータ側のアプリケーション1は、論理レコード221全体の受信がHCA2で完了した後、次の論理レコードの転送リクエストを開始することができる。



## 【0012】

図3では、HCA1側のアプリケーション1から転送リクエスト131によってRDMA転送を開始している。この場合、RDMA転送空間全体がひとつの論理レコードとなる。HCA2では、複数に分割されたパケットを受信しながら、必要に応じて再送要求を行う。すべてのパケット転送（論理レコード全体の転送）が正常に終了した時点で、HCA2側のアプリケーション2の完了キューに完了ステータス331を通知する。完了ステータス331を受けて、アプリケーション2は、転送された論理レコードであるRDMA転送空間を使った処理731を開始することができる。論理レコードの正常受信を完了したHCA2は、受信完了をHCA1に知らせ、HCA1はアプリケーション1（の完了キュー）に転送完了ステータス371を通知する。

## 【0013】

この例でも明らかなように、ターゲット側のアプリケーション2は、論理レコード（RDMA転送空間）全体の受信が完了してから完了ステータス331を受け取る。また、イニシエータ側のアプリケーション1は、論理レコード（RDMA転送空間）全体の受信がHCA2で完了した後、次の論理レコードの転送リクエストを開始することができる。

## 【0014】

このように従来のリライアブルデータ転送方法では、転送エラーを回避するための仕組みとして、上記2つの基本的な特徴をもつ必要があった。なお、特開平8-179999に開示されている従来のリライアブルデータ転送方法のように、エラー発生以前の転送データを保証する方式も知られているが、上記2つの特徴を満たす必要のあることに変わりはない。

## 【0015】

【特許文献1】特開平8-179999号公報

## 【0016】

【非特許文献1】InfiniBand Architecture Specification Release 1.0a

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0017】

従来のリライアブルデータ転送方法は、転送エラーを回避するために上記2つの特徴を利用しているのであるが、そのために、以下の課題をもつ。

まず、「イニシエータがターゲットでの論理レコード全体の正常受信完了を確認する必要がある」という特徴のため、イニシエータからターゲットへの論理レコード転送に要する時間以外に、ターゲットからイニシエータに転送完了を通知するための時間がオーバーヘッドとしてかかってしまう。このオーバーヘッド時間は、論理レコードのレコード長が短い時に顕著となり、通信路の効率を著しく低下させる。とくに、IOシステムの高機能化に伴ってターゲットでの処理タスクは増加しており、論理レコードの転送完了ステータス通知に要する時間は増大傾向にある。論理レコードの転送時間は、転送速度の向上にしたがって短縮されているので、イニシエータからの転送完了通知オーバーヘッドが相対的に増大して転送効率が悪化しており、その改善が課題となっている。

## 【0018】

次に、「ターゲットでは、受信した論理レコード全体にエラーの無いことを確認してから完了ステータスを通知する」という特徴のため、たとえ論理レコードの途中まで正常に受信ができたとしても、それをターゲットのアプリケーションで認識して有効に活用することができず、論理レコード全体の受信が完了するまで、受信論理レコードを利用したアプリケーション処理の開始を遅延させる必要がある。この処理開始までの遅延時間は、論理レコードのレコード長が長いときに顕著となり、アプリケーションの処理効率を低下させる。レコード長が長ければ、それだけ転送エラーの発生率が上昇し、転送エラーによる再送処理が行われると、上記遅延時間はさらに増加し、処理効率は低下する。大きなレコード長の転送や、転送エラーの発生時にも、アプリケーションの処理効率を低下させないことが、別の課題となっている。

**【0019】**

本発明の目的は、上記従来技術の欠点を改善し、高い転送効率と高いアプリケーション処理効率を同時に実現するデータ転送方法を提供することにある。より具体的には、ターゲットでの受信完了通知に要する時間とイニシエータへの転送完了通知に要する時間を実効的に削減するデータ転送方法およびかかるデータ転送方法を用いたディスク制御装置を提供することにある。

**【課題を解決するための手段】****【0020】**

上記目的を達成するため、本発明では、イニシエータからターゲットへのデータ転送の際、ターゲットで受信されたデータの通信エラーの有無をデータに付加されているエラーチェックコードを用いて確認し、通信エラーの有無をターゲットからイニシエータに転送ステータスとして返送し、転送ステータスによりデータ転送の際に転送エラーの発生が判明した場合、イニシエータからターゲットにデータの再送を行うリライアブル転送において、イニシエータとターゲット間のデータ転送単位である論理レコードの転送プロトコルにおいて、論理レコードを複数まとめて一括転送し、イニシエータでは、一括転送の単位で転送ステータスの確認を行い、イニシエータが発行した転送リクエストにより論理レコードの転送がなされ、ターゲットでは、転送リクエストに対応した完了ステータスを予め定められた一括転送条件に合致した論理レコードについて、その正常受信が完了した時点でターゲット内に存在する完了キューに通知するようにした。

**【0021】**

また、ターゲットでは、一括転送の途中で論理レコードの転送エラーをエラーチェックコードにより検出した際、一括転送の終了まで、当該論理レコードおよびそれ以降の論理レコードの受信を無視し、前記完了ステータスの通知を抑制するようにした。

**【0022】**

さらに、ターゲットでは、一括転送の途中で論理レコードの転送エラーをエラーチェックコードにより検出した際、一括転送の終了まで、当該論理レコード及び、それ以降の論理レコードのうち一括転送条件フィールドで指定されている論理レコードの受信を無視し、完了ステータスの通知を抑制するようにした。

**【0023】**

さらにまた、ターゲットでは転送エラーが検出された最初の論理レコードのIDを、一括転送の単位ごとに確認の行われる前記転送ステータスの中を含め、イニシエータでは転送ステータスをもとに、転送エラーの発生した論理レコードから再送を開始するようにした。

**【0024】**

また、一括転送の途中で、イニシエータ、もしくは、ターゲットからキャンセルリクエストを発行して、一括転送を中止することが可能であるようにした。

**【0025】**

さらに、ホストコンピュータとのインターフェースを有する複数のホストインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部とを有し、前記ホストインターフェース部は、前記ホストコンピュータに対しリード／ライトされるデータを一時的に格納するメモリを有し、前記メモリと前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリを有し、前記メモリと前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、複数のホストインターフェース部と複数のディスクインターフェース部との間の転送、複数のホストインターフェース部間の転送、あるいは 複数のホストインターフェース部もしくは複数のディスクインターフェース部と前記メモリ部との間の転送に上述のデータ転送方法を適用するようにした。

**【0026】**

さらにまた、ホストコンピュータとのインターフェースを有する複数のホストインター

フェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、複数のメモリ部と、複数のプロセッサ部とを有し、前記ホストインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記ホストコンピュータとの間でデータ転送を実行し、前記ディスクインターフェース部は、前記プロセッサ部と連携して前記メモリ部と前記磁気ディスク装置との間でデータ転送を実行するディスク制御装置において、前記複数のホストインターフェース部あるいは前記複数のディスクインターフェース部と前記プロセッサ部間の転送、複数のメモリ部とプロセッサ部間の転送、複数のメモリ部間の転送、あるいは、複数のプロセッサ部間の転送に上述のデータ転送方法を適用するようにした。

【発明の効果】

【0027】

以上説明したように、本発明のデータ転送方法では、一括転送中には、各論理レコードの転送リクエストが連続して処理され、イニシエータは、ターゲットからの個別の論理レコード転送完了ステータスを待たない。そのため、個別の論理レコード転送完了通知に要する時間が不要となり、転送効率を大幅に向上させることが可能となる。

【0028】

本発明によれば、一括転送条件に合致した正常受信の論理レコードについては、当該論理レコードの受信完了時点で完了ステータスがアプリケーションに通知される。そのため、多数の論理レコードを一括で転送した場合でも、一括転送全体が完了するのを待たずに、個別の論理レコードの受信が完了した時点で、対応するアプリケーションの処理を開始することができ、アプリケーションの処理効率を改善することが可能となる。とくに、一括転送中に受信エラーが発生した場合についても、エラー発生以前に正常受信した論理レコードについては、その受信完了時点で、アプリケーションの実行を開始できる。

【0029】

本発明によれば、一括転送における一括転送条件をさらに詳細に規定することができ、一括転送中に転送エラーが生じた場合でも、データ依存関係がなく、到着順序で問題を生じる可能性の無い論理レコードについては、受信動作を継続することが可能となる。そのため、アプリケーションの処理効率を高めることができ、再送レコード数を減らすことでより効率のよい転送を実現することができる。

【0030】

本発明に拠れば、再送の必要な論理レコードについてのみ再送処理を行うことで、再送処理に要する時間を最小限に抑えることができる。したがって、再送レコード数を減らすことでより効率のよい転送を実現することができる。

【0031】

本発明に拠れば、一括転送の途中で転送エラー発生などにより、その後の一括転送の継続が不要になった場合、当該一括転送を途中で中止することが可能となる。そのため、不要な論理レコードの転送を抑制することで、より効率のよい転送を実現することができる。

【0032】

また、本発明のディスク制御装置では、ホストコンピュータ部とディスクインターフェース部の転送効率を大幅に改善することができる。特に、ディスク制御装置内ネットワークの転送効率の向上と、ホストインターフェース部、および、ディスクインターフェース部からのI/Oリクエストに対するレスポンス時間の短縮が可能となり、ディスク制御装置の処理性能を高めることができる効果を奏する。

【発明を実施するための最良の形態】

【0033】

以下、本発明の実施例を図面を用いて詳述する。

《実施例1》

図1及び図8に、本発明の一実施例を示す。本発明のデータ転送方法においては、複数の論理レコードをまとめて一括転送を行う。そのため、図8に示すように、イニシエータ

の送信キューに複数の転送リクエストが連続して格納される。各転送リクエストには、論理レコードの転送ごとに一意に定まる転送IDと、転送リクエストの動作を規定する転送命令コードと、転送レコードバッファの先頭アドレスおよびそのバッファ長を含んでいる。図8では、転送リクエスト101がレコードバッファ89の単純転送、転送リクエスト105がRDMA転送領域87のRDMA転送を示しており、これらが一括転送リクエストとして定義されている。

#### 【0034】

また、図8には、論理レコードの受信完了を通知するために用いられるターゲットの完了キュー39とそこに格納される(論理レコードの)完了ステータス300の構成も示してある。完了ステータス300は、論理レコードの転送IDと対応するステータスIDと、その完了ステータスコードが含まれている。

#### 【0035】

この一括転送リクエストが処理される様子の一例を図1に示した。ここでは、アプリケーション1がHCA1とHCA2を介してアプリケーション2に転送リクエスト101-105を発行している。HCA1では、転送リクエスト101によって、論理レコード201の転送を開始する。HCA2では、論理レコード201の正常受信完了を確認した後、アプリケーション2の完了キューに対して完了ステータス301を通知している。アプリケーション2では、完了ステータス301を受けて、論理レコード201に対する処理動作701を開始することができる。

#### 【0036】

本発明の高速データ転送方法では、従来方式と異なって、論理レコード102の転送を開始するのに、HCA2からの論理レコード101の転送完了ステータス通知を待たない。HCA1では論理レコード101の転送が終了した後、ただちに論理レコード102の転送を開始する。また、HCA2では、受信した論理レコード102にエラーが含まれていた場合、従来方式のように論理レコード102の再送要求処理を行うのではなく、一括転送完了ステータスに論理レコード102の受信エラーを記録した後、次の論理レコード103の処理に移る。一旦、転送リクエスト105までのすべての一括転送リクエストが処理された時点で、HCA2はまとめて、一括転送完了ステータス351をHCA1に返送する。これを受けて、HCA1では、HCA2で正常受信の完了していない論理レコードの再送を行う。HCA2では、再送によって正常に受信できたレコード212-215については、完了キューへの完了ステータス312-315の通知を行い、アプリケーション2では対応する処理712-715を開始する。一括転送完了ステータス352によって、すべての一括転送リクエストの正常転送完了ステータスが通知されることで、一連の一括転送が終了する。

#### 【0037】

本実施例によれば、一括転送中には、各論理レコードの転送リクエストが連続して処理され、イニシエータは、ターゲットからの個別の論理レコード転送完了ステータス通知を待たない。そのため、個別の論理レコード転送完了通知に要する時間が不要となり、転送効率を大幅に向上させることが可能となる。

また、本実施例によれば、一括転送条件に合致した正常受信の論理レコードについては、当該論理レコードの受信完了時点で完了ステータスがアプリケーションに通知される。そのため、多数の論理レコードを一括で転送した場合でも、一括転送全体の完了を待たずに、個別の論理レコード受信の完了した時点で、対応するアプリケーション処理を開始することができ、アプリケーション処理効率を改善することが可能となる。とくに、一括転送中に転送エラーの発生した場合についても、エラー発生以前に正常受信した論理レコードについては、その受信完了時点で、アプリケーションの実行を開始することができる。

#### 【0038】

##### 《実施例2》

本発明のデータ転送方法では、一括転送の論理レコードを受信する際、一括転送条件に合致したものだけを選択的に受信する。これにより、一括転送内の論理レコード間で依存

関係があっても、受信レコードの時系列順序を保証することが可能となる。例えば、ある論理レコードに受信エラーが含まれていた場合、一括転送中のそれ以降の論理レコードを無視してしまう方法を実施することができる。この方法により、すべての一括転送リクエストを連続して処理した後に、まとめて再送処理をおこなう本発明の方式では、受信論理レコードの依存関係に問題を生じることがなく、従来方式のようにイニシエータでターゲットからの転送完了通知を待つ必要がないという利点がある。

#### 【0039】

##### 《実施例3》

本発明のデータ転送方法で用いられる一括転送条件について、図9を用いてさらに詳しい説明をおこなう。図9は、本発明のデータ転送方法で使用されるパケットの構造を示したものである。ルーティングに関する情報をもったルーディングヘッダ441と、トランスポート処理に関する情報の入ったトランスポートヘッダ442と、論理レコード情報の入るペイロード443とエラーチェックコードであるCRC444とからなる。ルーディングヘッダは、イニシエータ・ターゲットの宛先アドレスと、パケットの優先度情報、パケット長を含んでいる。トランスポートヘッダには、転送処理を規定する処理動作コードと宛先のキューペア番号、パケット順序番号と、一括転送の動作を規定する一括転送フラグ450、一括転送条件フィールド451が含まれている。

#### 【0040】

一括転送フラグ450は、当該パケットが一括転送されていることを示すものであり、このフラグを確認することで、一括転送中なのかどうかを判断することができる。一括転送条件フィールド451は、一括転送中の各論理レコードのデータ依存関係を示したものである。各論理レコードに対応するビットフィールドに1がセットされると、その論理レコードが他の論理レコードと依存関係の無いことを示す。一括転送中にエラーが発生した場合でも、この一括転送条件フィールドで1のセットされている論理レコードについては、受信処理を実行してもデータの到着順序で問題の生じることが無い。

#### 【0041】

また、後に明らかになるように、一括転送フラグの情報や、一括転送条件フィールドの情報をペイロードの中にいれて、あらかじめイニシエータからターゲットに通知しておいてもよい。その場合、各パケット中のトランスポートヘッダに、一括転送フラグや、一括転送条件フィールドがなくてもよい。

#### 【0042】

本実施例によれば、一括転送における一括転送条件をさらに詳細に規定することができる。したがって、一括転送中に転送エラーが生じた場合でも、データ依存関係がなく、到着順序で問題を生じる可能性の無い論理レコードについては、受信動作を継続することが可能となる。そのため、アプリケーションの処理効率を高めることができ、再送レコード数を減らすことでより効率のよい転送を実現することができる。

#### 【0043】

##### 《実施例4》

本発明のデータ転送方法の動作フローを、図10を用いて説明する。ここでは一括転送を開始するにあたり、イニシエータがターゲットに一括転送モード開始リクエストを発行している。これは、上述のように、一括転送条件フィールドなどの一括転送に関する情報を、通常のリライアブル送信パケットのペイロードに含めて送信するものである。ターゲットから当該パケットに対する受信完了通知が届いたら、イニシエータ・ターゲットとも一括転送モードになる。ここで、イニシエータからは、送信キューに格納されている一括転送リクエストの処理を順次開始し、ターゲットでは、それに対応したレスポンス処理を行う。一連の一括転送リクエストの処理が終了した時点で、ターゲットから一括転送モードの終了リクエストを、通常のリライアブル送信パケットとして送信する。

#### 【0044】

このとき、ターゲットでの一括転送受信記録である一括転送完了ステータスをペイロードの一部に含めて送信する。イニシエータからは、その受信完了通知を返信し、イニシエ

ータ・ターゲットは一括転送モードを解除する。イニシエータと、ターゲットは、一括転送完了ステータスを参照して、必要ならば、再送処理を開始する。すなわち、再送すべき論理レコードが存在すれば、それを順次、上記の手順に従い、一括転送する。一括転送リクエストのすべての論理レコードがターゲットに正常受信されたなら、一括転送を終了する。

#### 【0045】

本実施例において、ターゲットでは、転送エラーが発生した以降の当該一括転送モード中の論理レコード受信については無視し、アプリケーションの完了キューにも通知しない。そして、上記の一括転送完了ステータスの中に、一括転送モード中で転送エラーの検出された最初の論理レコードの転送IDを含めて返送する。イニシエータでは、一括転送完了ステータスに含まれている、エラーの検出された論理レコードから再送を開始する。

#### 【0046】

本実施例に拠れば、一括転送中で転送エラーの生じるまでの正常受信が完了した論理レコードに関しては、受信完了後ただちに、対応するアプリケーションの処理を開始することができる。したがって、転送エラーが発生した際にも、アプリケーションの処理効率を高めることができ、再送レコード数を減らすことでより効率のよい転送を実現することができる。

#### 【0047】

##### 《実施例5》

別の実施例では、上記の一括転送完了ステータスの中に、再送の必要な論理レコードの転送IDリストを含めて返送する。イニシエータでは、ターゲットから通知された前記転送IDリストを参照して再送処理を行う。

#### 【0048】

本実施例に拠れば、再送の必要な論理レコードについてのみ再送処理を行うことで、再送処理に要する時間を最小限に抑えることができる。したがって、再送レコード数を減らすことでより効率のよい転送を実現することができる。

#### 【0049】

##### 《実施例6》

本発明のデータ転送方法における動作の詳細について、図11、図12を用いて説明を行う。図11は、図10内の一括転送モードにおけるSENDリクエストの動作フローを示したものである。当該論理レコードが一括転送条件に合致していれば、送信処理1201を行い、正常に送信できたかどうか、送信ステータス記録処理1221を行う。図12は、SENDリクエストに対応したレスポンスの動作フローを示している。当該論理レコードが一括転送条件に合致していれば、受信処理1101を行い、正常に受信できた場合完了キュー通知1110を行って、最後に受信ステータス記録処理1121を行う。

#### 【0050】

SENDリクエスト処理、SENDレスポンス処理とも、一括転送条件に合致しない論理レコードについては、当該レコードをスキップする。その際、相互にキャンセルリクエスト1112および1212を発行して、当該一括転送を途中で中止することが可能である。キャンセルリクエストを受けたほうは、その応答1111もしくは1211を返した後、一括転送モードを解除する。キャンセルリクエストを出したほうは、相手からその応答が戻ってきた後、一括転送モードを解除する。

#### 【0051】

本実施例に拠れば、一括転送の途中で転送エラー発生などにより、その後の一括転送の継続が不要になった場合、当該一括転送を途中で中止することが可能となる。そのため、不要な論理レコードの転送を抑制することで、より効率のよい転送を実現することができる。

#### 【0052】

##### 《実施例7》

図13、図14、図15に本発明のディスク制御装置に関する一実施例を示す。図13

において、ディスク制御装置500は、ホストコンピュータ560とホストインターフェースネットワーク501で接続している複数のホストインターフェース部510と、磁気ディスク装置570とディスクインターフェースネットワーク502で接続している複数のディスクインターフェース部520とを有しており、複数のホストインターフェース部510と複数のディスクインターフェース部520は、ディスク制御装置内ネットワーク503で接続されている。

#### 【0053】

図14にホストインターフェース部510の構成を示す。複数のホストチャンネルインターフェース511とプロセッサ512とメモリ513とHCA603を有し、それらがホストハブ514を介して接続されている。

#### 【0054】

図15にディスクインターフェース部520の構成を示す。複数のディスクチャンネルインターフェース521とプロセッサ522とメモリ523とHCA604を有し、それらがディスクハブ524を介して接続されている。また、ディスクハブ524には、キャッシュメモリ525も接続している。

#### 【0055】

各ホストインターフェース部510は、ホストコンピュータ560とのインターフェースとキャッシュメモリ525との間のデータ転送を実行し、各ディスクインターフェース部520は、磁気ディスク装置570とのインターフェースとキャッシュメモリ525との間のデータ転送を実行する。

#### 【0056】

ホストインターフェース部510と、ディスクインターフェース部520は、HCA603とHCA604を介してデータ転送を行う。その際、ホストインターフェース部は、複数のホストチャンネルからのコマンドやデータをホストハブでまとめて、ディスクインターフェース部に転送する。その際に、上述のデータ転送方法を用いる。なお、HCAは、図5に示した機能と同等の機能を有するものであればよく、例えば、その一部の機能をプロセッサ512やプロセッサ522のソフトウェア処理により実現してもかまわない。

#### 【0057】

本実施例によると、ホストコンピュータ部とディスクインターフェース部の転送効率を大幅に改善することができる。特に、ディスク制御装置内ネットワーク503の転送効率の向上と、ホストインターフェース部、および、ディスクインターフェース部からのI/Oリクエストに対するレスポンス時間の短縮が可能となり、ディスク制御装置の処理性能を高めることができる。

#### 【0058】

##### 《実施例8》

図13に示したディスク制御装置は、その信頼性を向上させるため、複数のディスクインターフェース部に内蔵されているキャッシュメモリに、冗長にデータを格納している。そのため、ホストインターフェース部510から、ディスクインターフェース部520に書き込み要求があった場合、あるディスクインターフェース部内のキャッシュメモリ525にデータを格納した後、別のディスクインターフェース部内のキャッシュメモリに対しても、冗長書き込みを実行する。この複数のディスクインターフェース部間での冗長書き込みに、上述のデータ転送方法を使用する。

#### 【0059】

冗長書き込みは、ディスク制御装置の信頼性向上のために不可欠ではあるが、それによるディスク制御装置内ネットワークへの負荷増加とそれに起因するシステム性能劣化が課題であった。本実施例により、ディスク制御装置内ネットワークの転送効率を向上させて、ディスク制御装置の処理性能を高めることが可能となる。

#### 【0060】

##### 《実施例9》

図13に示したディスク制御装置では、仮想化機能実現などのために他のディスク制御

装置と連携して動作する場合に、ホストインターフェース部間でデータ転送を行う必要がある。このホストインターフェース部間のデータ転送に、上述のデータ転送方法を適用する。

#### 【0061】

ホストインターフェース部間転送の場合、複数のディスク制御装置を経由してのデータアクセスとなるので、そのレスポンス時間を、可能な限り短くする必要がある。本実施例によるとホストインターフェース部間でのレスポンス時間の大幅な短縮が可能となり、ディスク制御装置の処理性能が向上する。

#### 【0062】

##### 《実施例10》

図16に示したディスク制御装置では、ディスク制御装置500が、ホストコンピュータ560とホストインターフェースネットワーク501で接続している複数のホストインターフェース部610と、磁気ディスク装置570とディスクインターフェースネットワーク502で接続している複数のディスクインターフェース部620と、複数のメモリ部580と、複数のプロセッサ部590を有しており、複数のホストインターフェース部610と複数のディスクインターフェース部620と、複数のメモリ部580と、複数のプロセッサ部590とは、ディスク制御装置内ネットワーク503で接続されている。

#### 【0063】

前記複数のホストインターフェース部501は、上記複数のプロセッサ590と連携して上記メモリ部580とデータ転送を実行し、あるいは複数のディスクインターフェース部620は上記プロセッサ部590と連携して上記磁気ディスク装置570と上記メモリ部580との間でデータ転送を実行する。このデータ転送に上述の本発明のデータ転送方法を用いる。

#### 【0064】

本実施例によると、ディスク制御装置内ネットワーク503の転送効率の向上と、ホストインターフェース部501、および、ディスクインターフェース部620からのI/Oリクエストに対するレスポンス時間の短縮が可能となり、ディスク制御装置の処理性能を高めることができる。

#### 【0065】

また、上記複数のメモリ部580とプロセッサ部590間のデータ転送や上記複数のメモリ部580間、あるいは、プロセッサ部590間のデータ転送にも上述の本発明の高速データ転送方法を用いることができ、同様にディスク制御装置の処理性能を高めることができる。

#### 【図面の簡単な説明】

#### 【0066】

【図1】本発明によるデータ転送方法の動作原理を示す図である。

【図2】従来のデータ転送方法の動作原理を示す図である。

【図3】従来のデータ転送方法の動作原理を示す図である。

【図4】IOシステムの構成を示す図である。

【図5】ホストチャネルアダプタ（HCA）の構成を示す図である。

【図6】従来の単純データ転送方法の動作詳細を示す図である。

【図7】従来のRDMAデータ転送方法の動作詳細を示す図である。

【図8】本発明のデータ転送方法で使用する転送リクエストと論理レコードの構造を示す図である。

【図9】本発明のデータ転送方法で使用するパケット構造を示す図である。

【図10】本発明のデータ転送方法の動作フローを示す図である。

【図11】本発明のデータ転送方法におけるSENDリクエストの動作フローを示す図である。

【図12】本発明のデータ転送方法におけるSENDレスポンスの動作フローを示す図である。



【図 13】本発明のディスク制御装置の構成を示す図である。

【図 14】本発明のディスク制御装置で使用されるホストインターフェース部の構成を示す図である。

【図 15】本発明のディスク制御装置で使用されるディスクインターフェース部の構成を示す図である。

【図 16】本発明のディスク制御装置の構成を示す図である。

【図 17】本発明のディスク制御装置で使用されるホストインターフェース部の構成を示す図である。

【図 18】本発明のディスク制御装置で使用されるディスクインターフェース部の構成を示す図である。

【符号の説明】

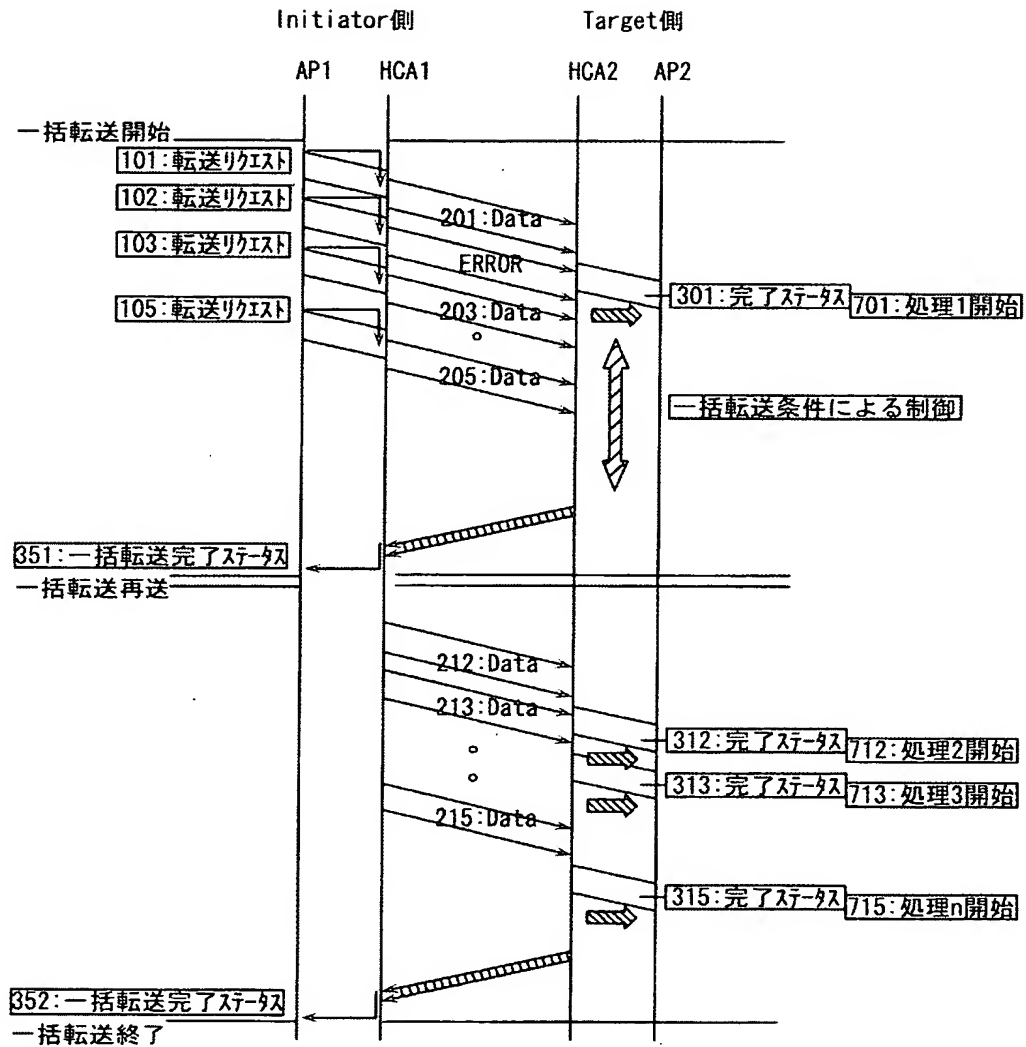
【0067】

11、12、13、14、15、16、17、18、19…送信キュー、  
21、22、23、24、25、26、27、28…受信キュー、  
31、32、33、34、35、36、39…完了キュー、  
41、42、43、44、45、46、47、48…キューペア、  
51、52、53、54…プロセス、  
61、62…プロセスコード、  
71、72、73、74…プロセスバッファ、  
75、76…アプリケーションメモリ空間、  
81、82、83、84、89…レコードバッファ、  
85、86、87…RDMA転送空間、  
100、101、102、103、105、121、125、131…転送リクエスト  
、  
201、205、212、213、215、221、225、231、235、232  
、234…転送データ、  
300、301、312、313、315、321、325、331…(論理レコード)  
完了ステータス、  
351、352…一括転送完了ステータス、  
361、365、371…転送完了ステータス、  
401、402、403、404、411、413、414、421、422、423  
、424、431、433、434、440…パケット、  
441…ルーティングヘッダ、  
442…トランスポートヘッダ  
443…ペイロード、  
444…CRC、  
450…一括転送フラグ、  
451、456…一括転送条件フィールド、  
455…一括転送レコード数、  
500…ディスク制御装置、  
501…ホストインターフェースネットワーク、  
502…ディスクインターフェースネットワーク、  
503…ディスク制御装置内ネットワーク、  
510…ホストインターフェース部、  
511…ホストチャネルインターフェース、  
512…プロセッサ、  
513…メモリ、  
514…ホストハブ、  
520…ディスクインターフェース部、  
521…ディスクチャネルインターフェース、

5 2 2…プロセッサ、  
5 2 3…メモリ、  
5 2 4…ディスクハブ、  
5 2 5…キャッシュメモリ  
5 6 0…ホストコンピュータ、  
5 7 0…磁気ディスク装置、  
5 8 0…メモリ、  
5 9 0…プロセッサ、  
6 0 1、6 0 2、6 0 3、6 0 4…ホストチャネルアダプタ、  
6 1 0…ホストインターフェース部、  
6 1 1、6 1 2、6 1 3…受信ポート、  
6 2 0…プロセッサ部、  
6 2 1、6 2 2、6 2 3…送信ポート、  
6 3 1…受信リンク論理、  
6 3 2…受信トランスポート論理、  
6 3 3…受信プロセッサ、  
6 4 1…送信リンク論理、  
6 4 2…送信トランスポート論理、  
6 4 3…送信プロセッサ、  
6 5 0…メモリ、  
6 6 0…接続インターフェース、  
6 7 0…内部バス  
7 0 1、7 1 2、7 1 3、7 1 5、7 2 1、7 2 5、7 3 1…処理開始、  
1 0 0 1…一括送信モード開始リクエスト、  
1 0 0 2…一括転送モード開始レスポンス、  
1 0 1 1…送信リクエスト、  
1 0 1 3…R DMA 送信リクエスト、  
1 0 1 5…送信レスポンス、  
1 0 1 7…R DMA 送信レスポンス、  
1 0 2 1…一括転送モード終了レスポンス、  
1 0 2 2…一括転送モード終了リクエスト  
1 1 0 1…受信、  
1 1 1 0…完了通知、  
1 1 1 1…キャンセルレスポンス、  
1 1 1 2…キャンセルリクエスト、  
1 1 2 1…受信ステータス記録、  
1 2 0 1…送信、  
1 2 1 1…キャンセルレスポンス、  
1 2 1 2…キャンセルリクエスト、  
1 2 2 1…送信ステータス記録。

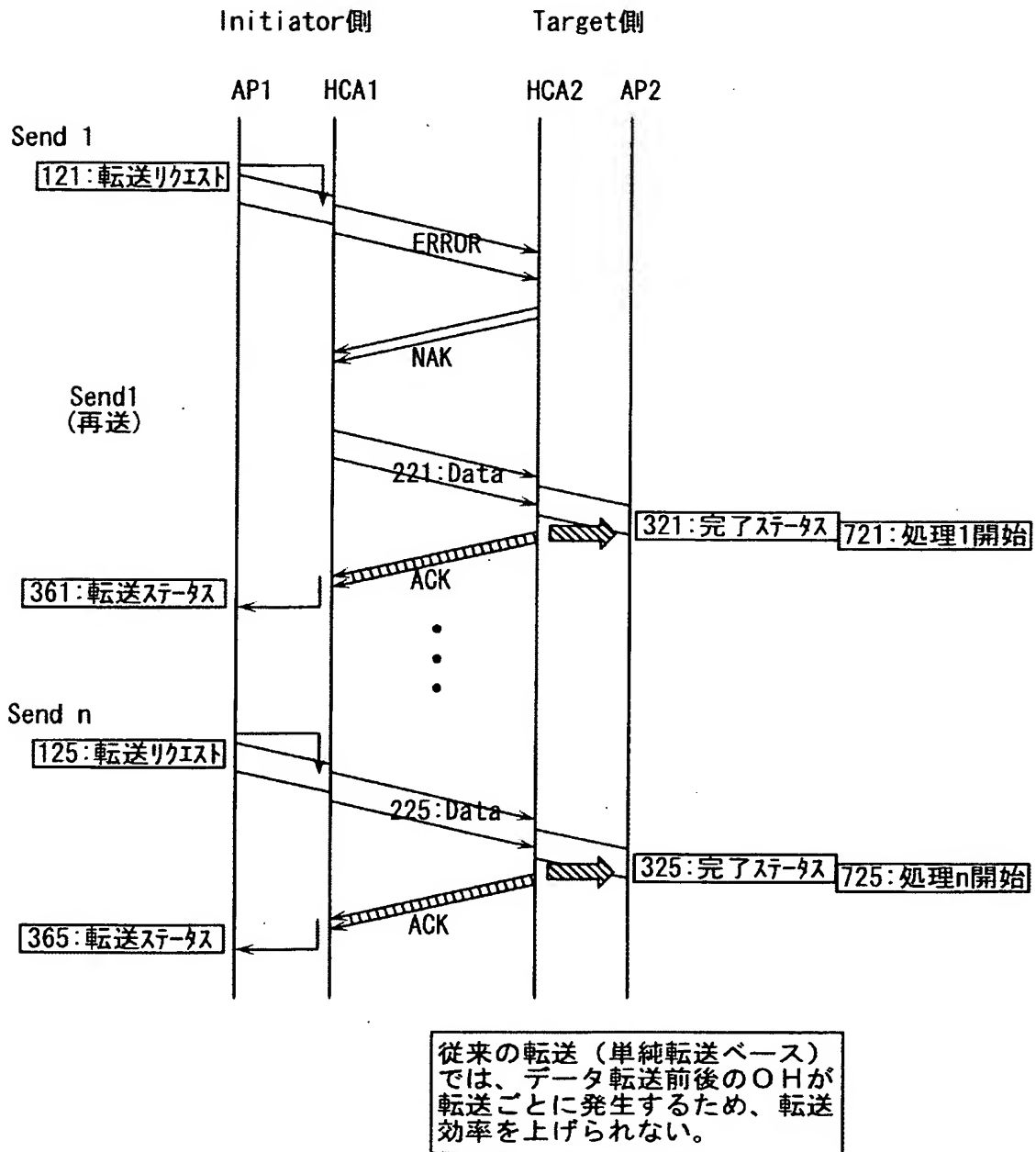
【書類名】 図面  
【図 1】

図 1



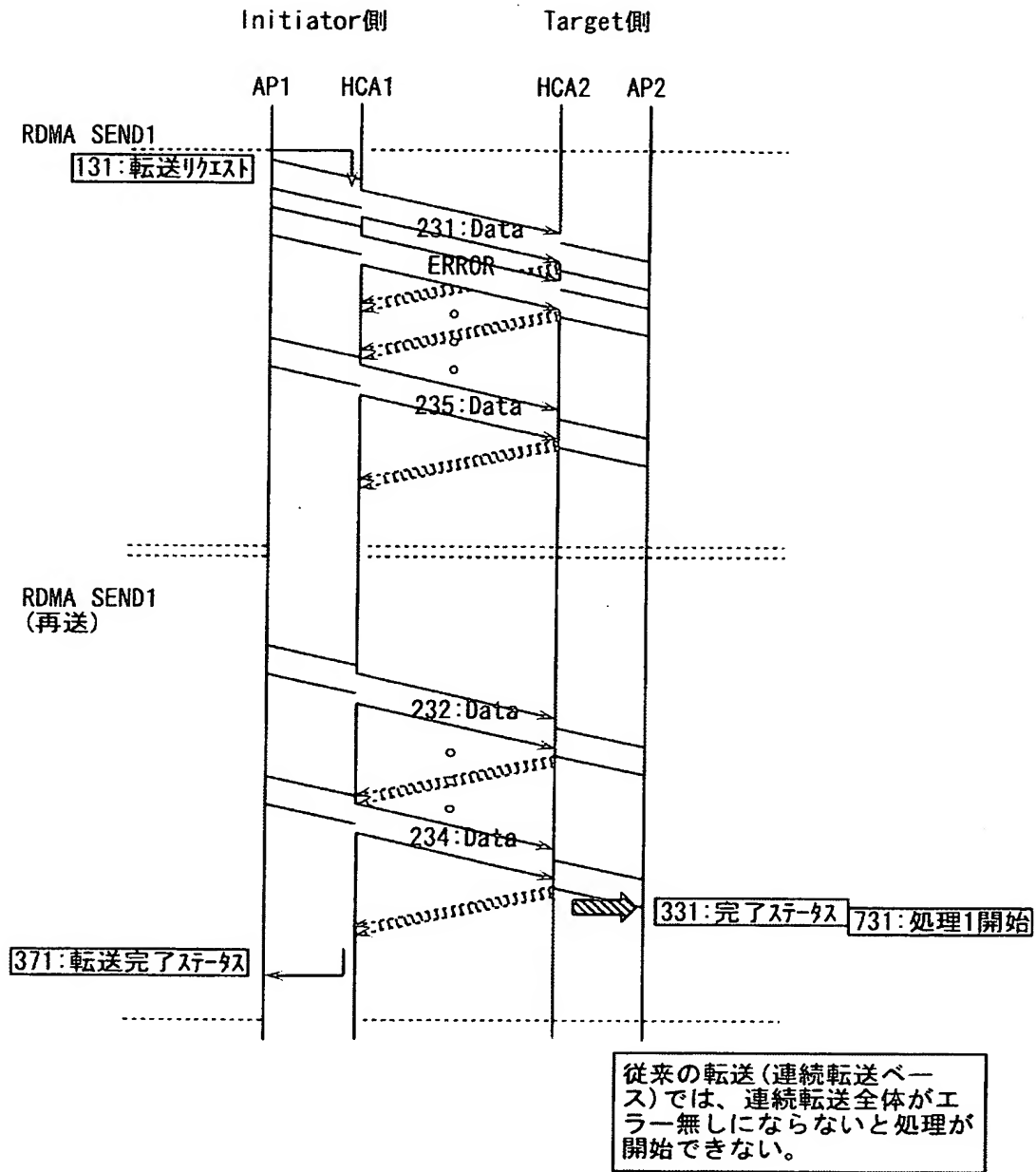
【図 2】

図 2



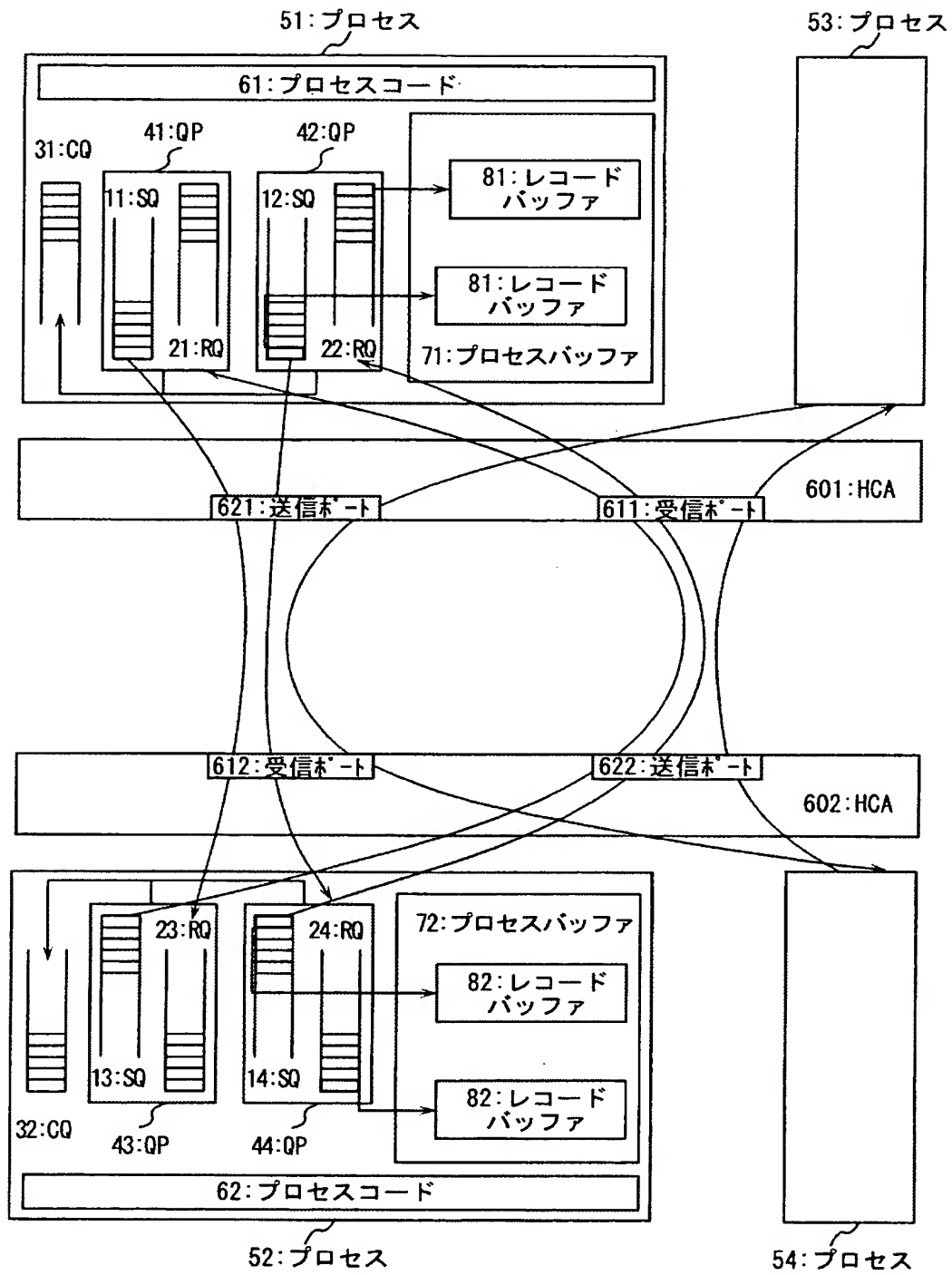
【図 3】

図 3



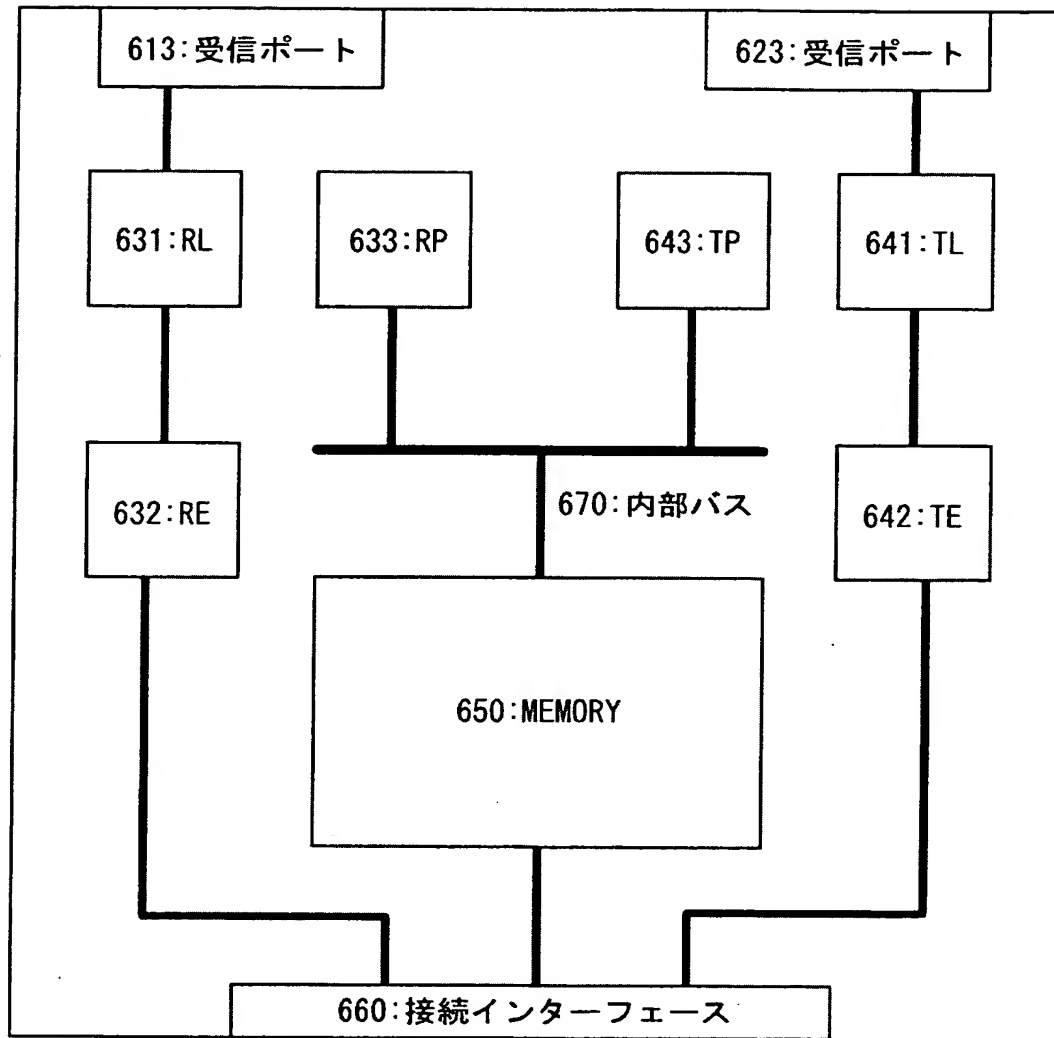
【図 4】

図 4



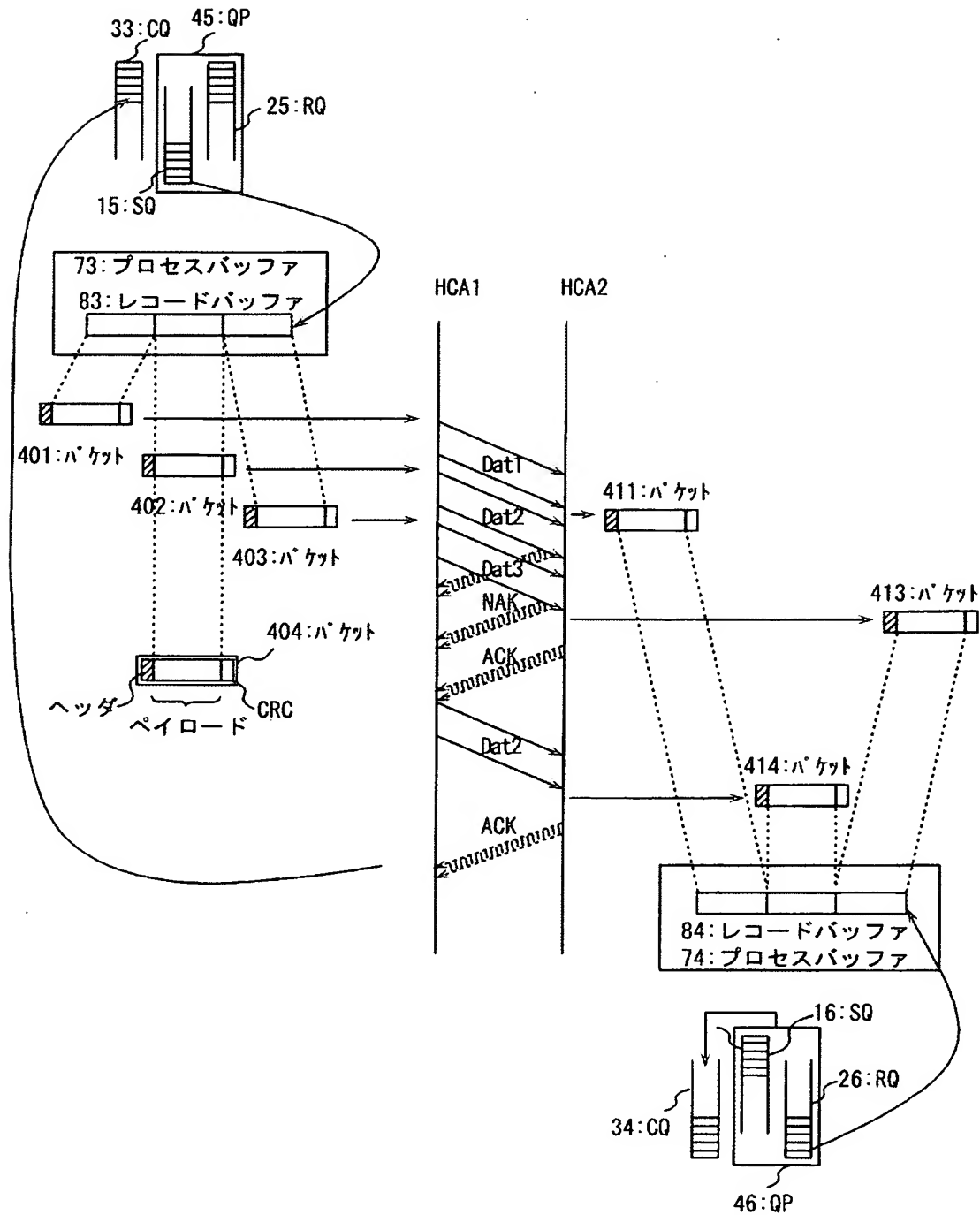
【図 5】

図 5



【図 6】

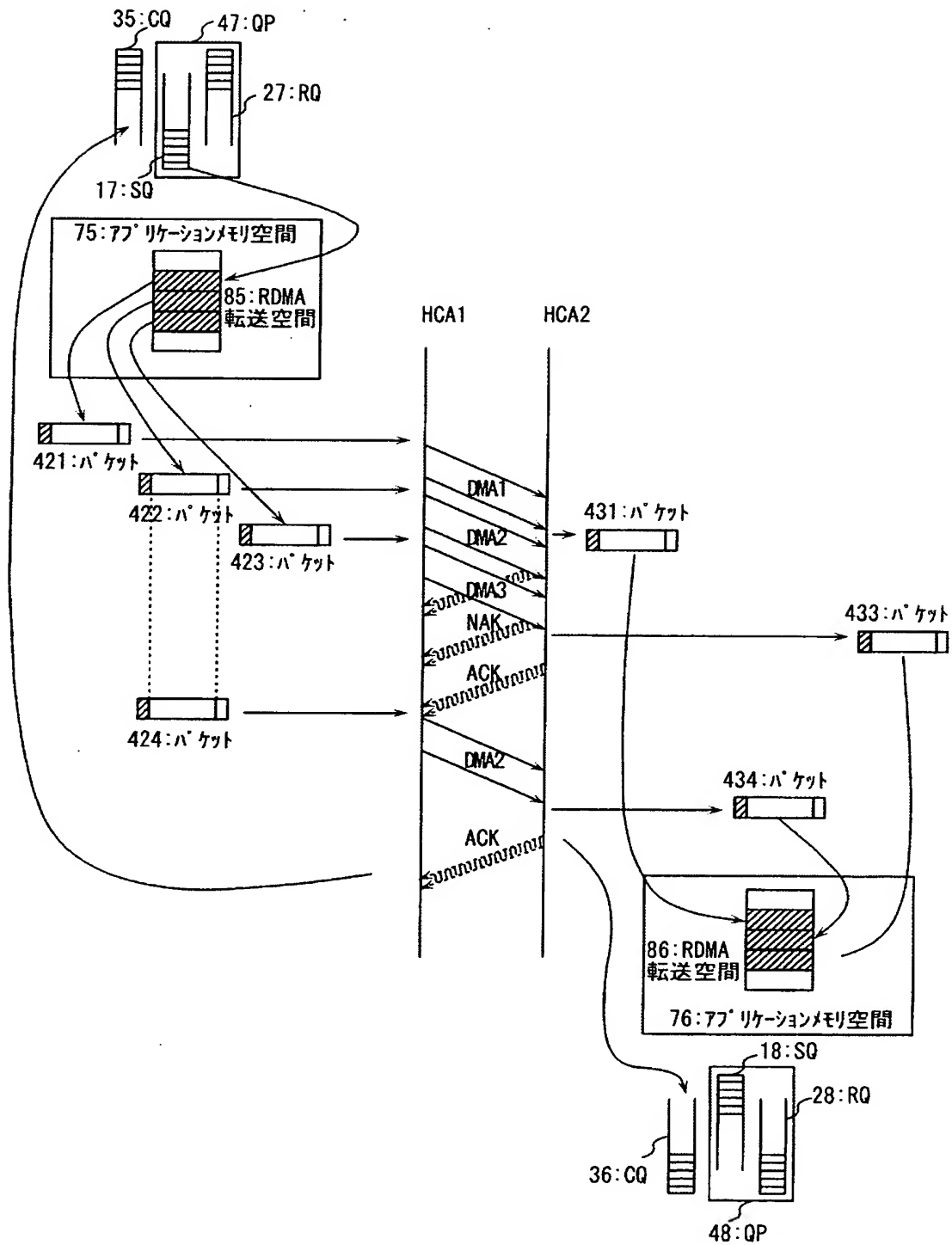
図 6





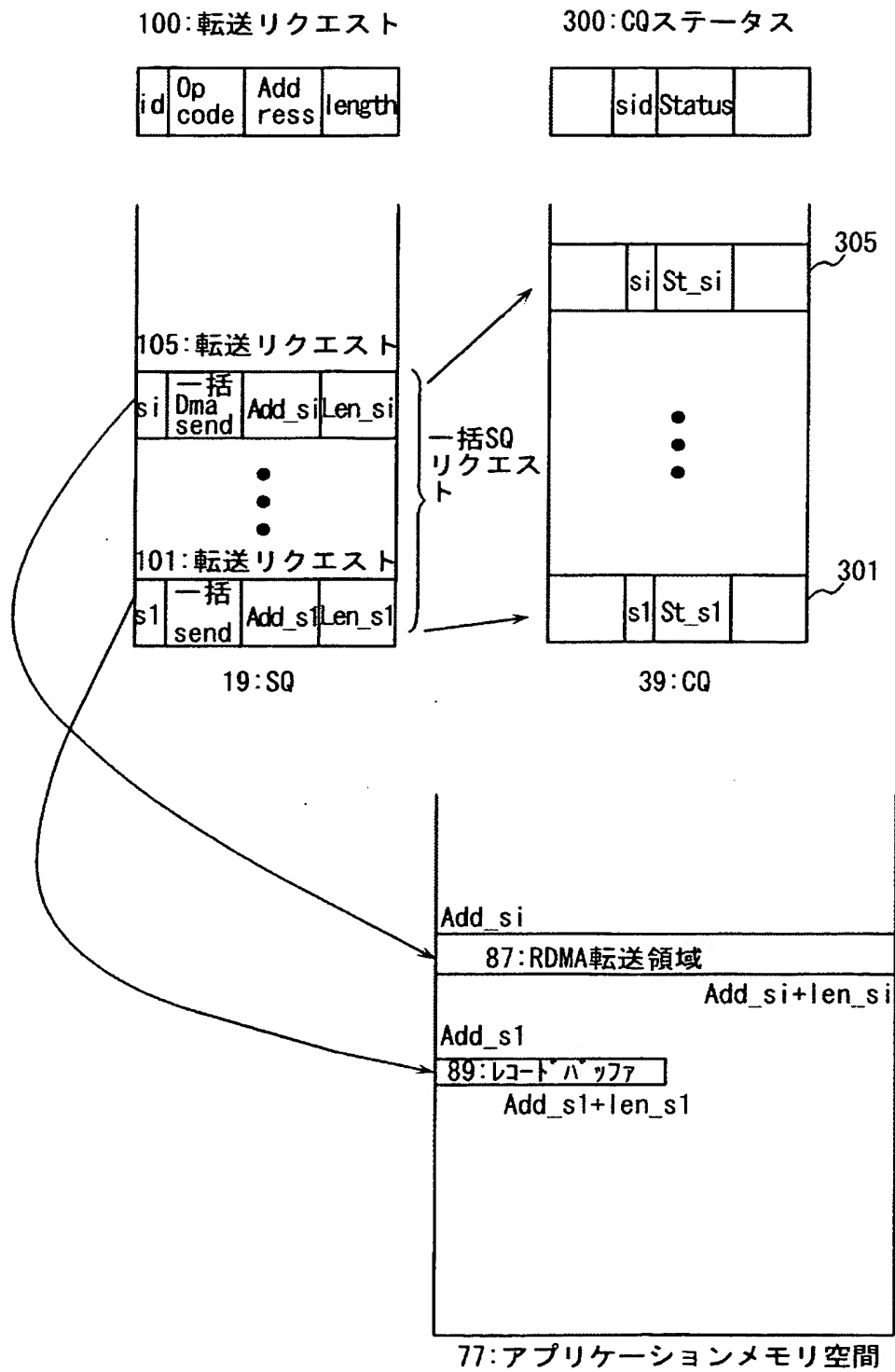
【図 7】

図 7



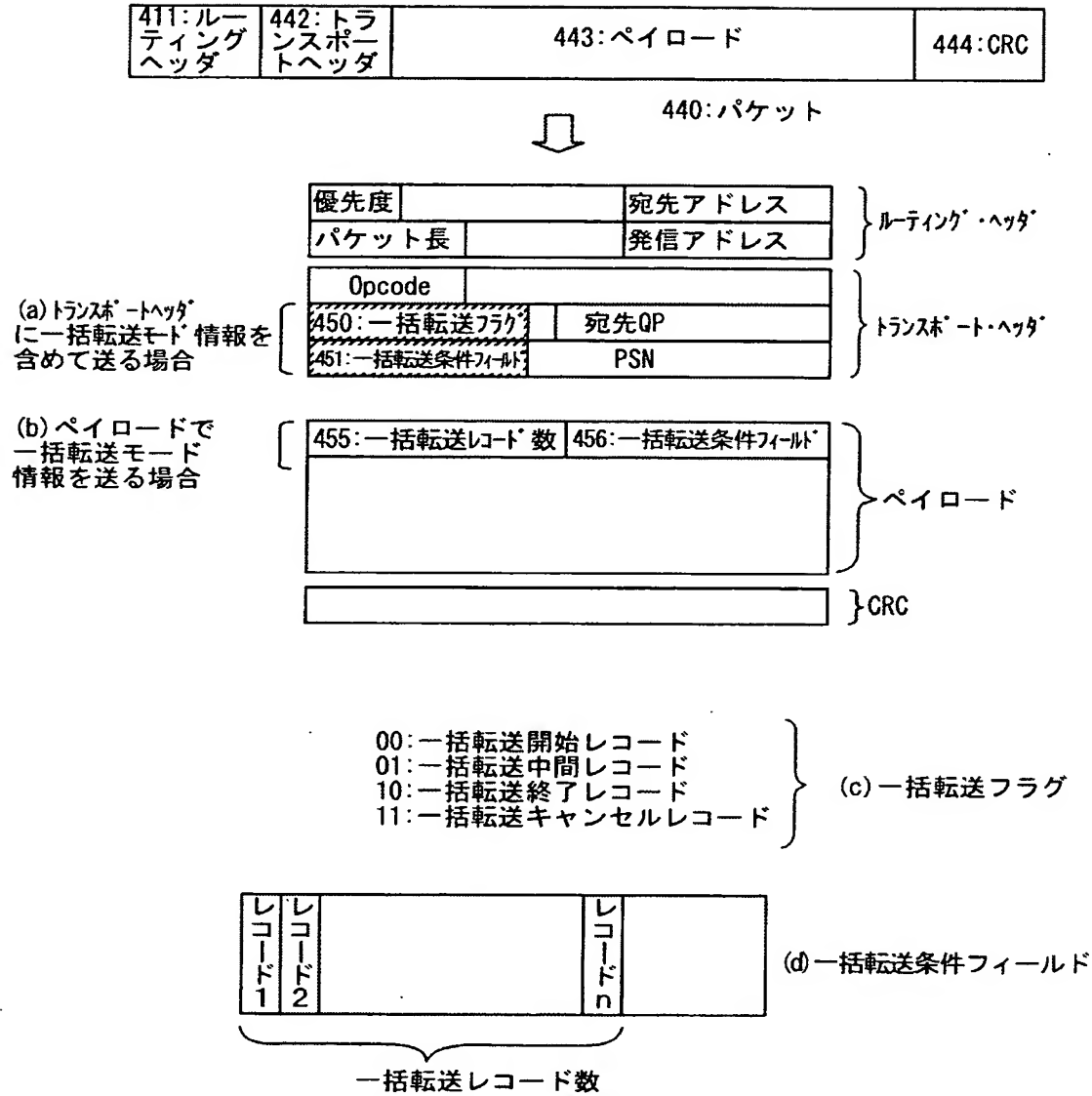
【図 8】

図 8



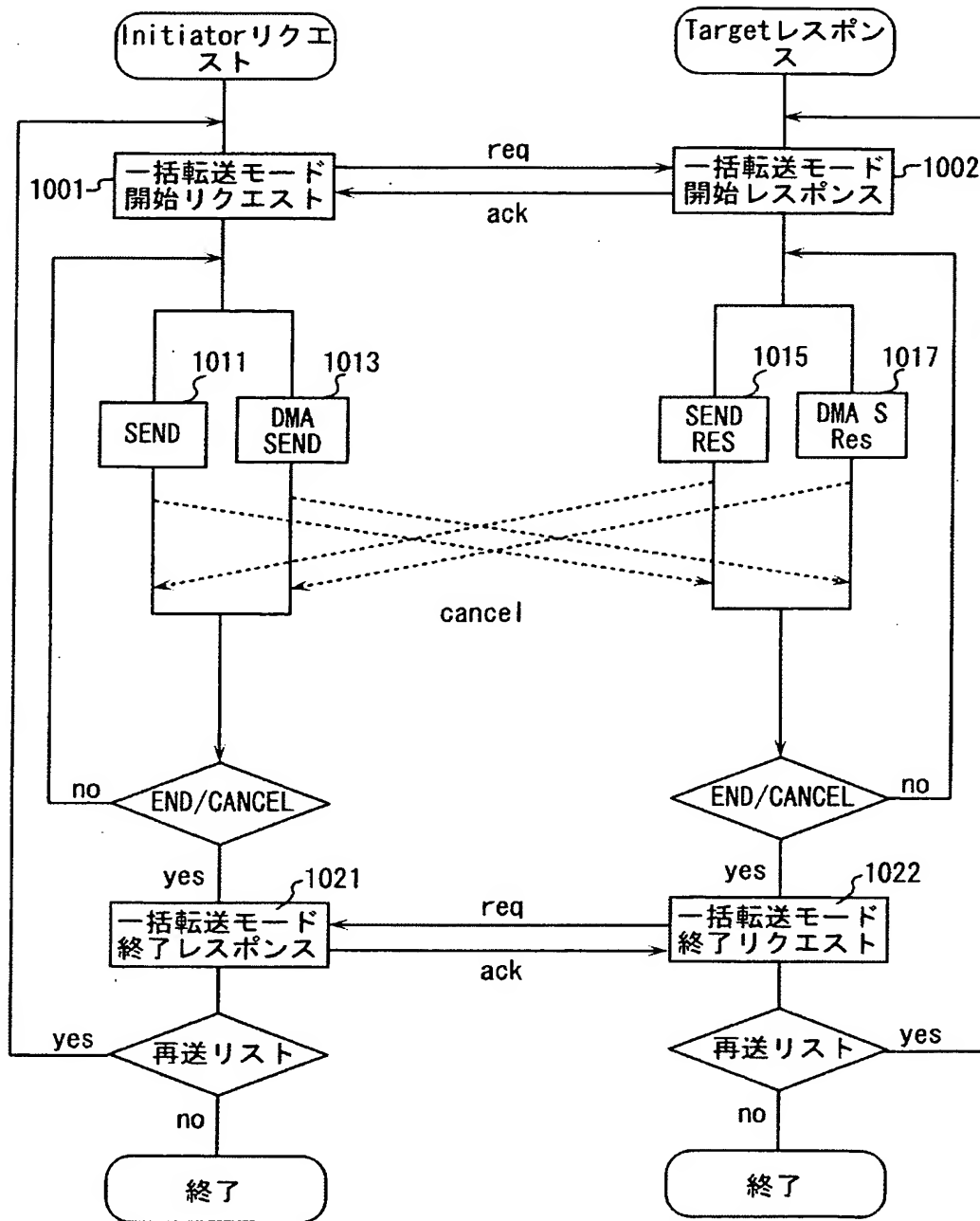
【図 9】

図 9



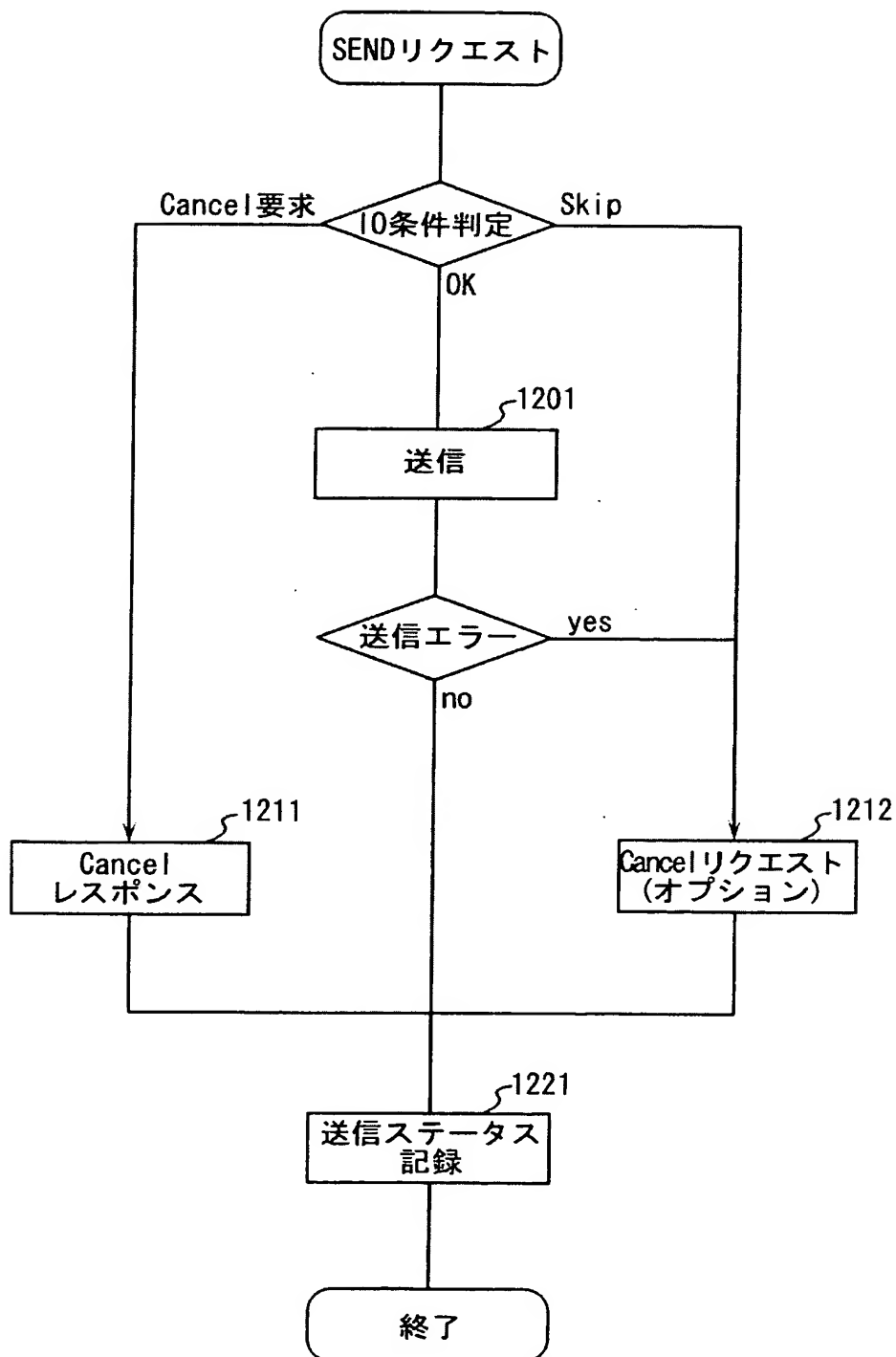
【図 10】

図 10



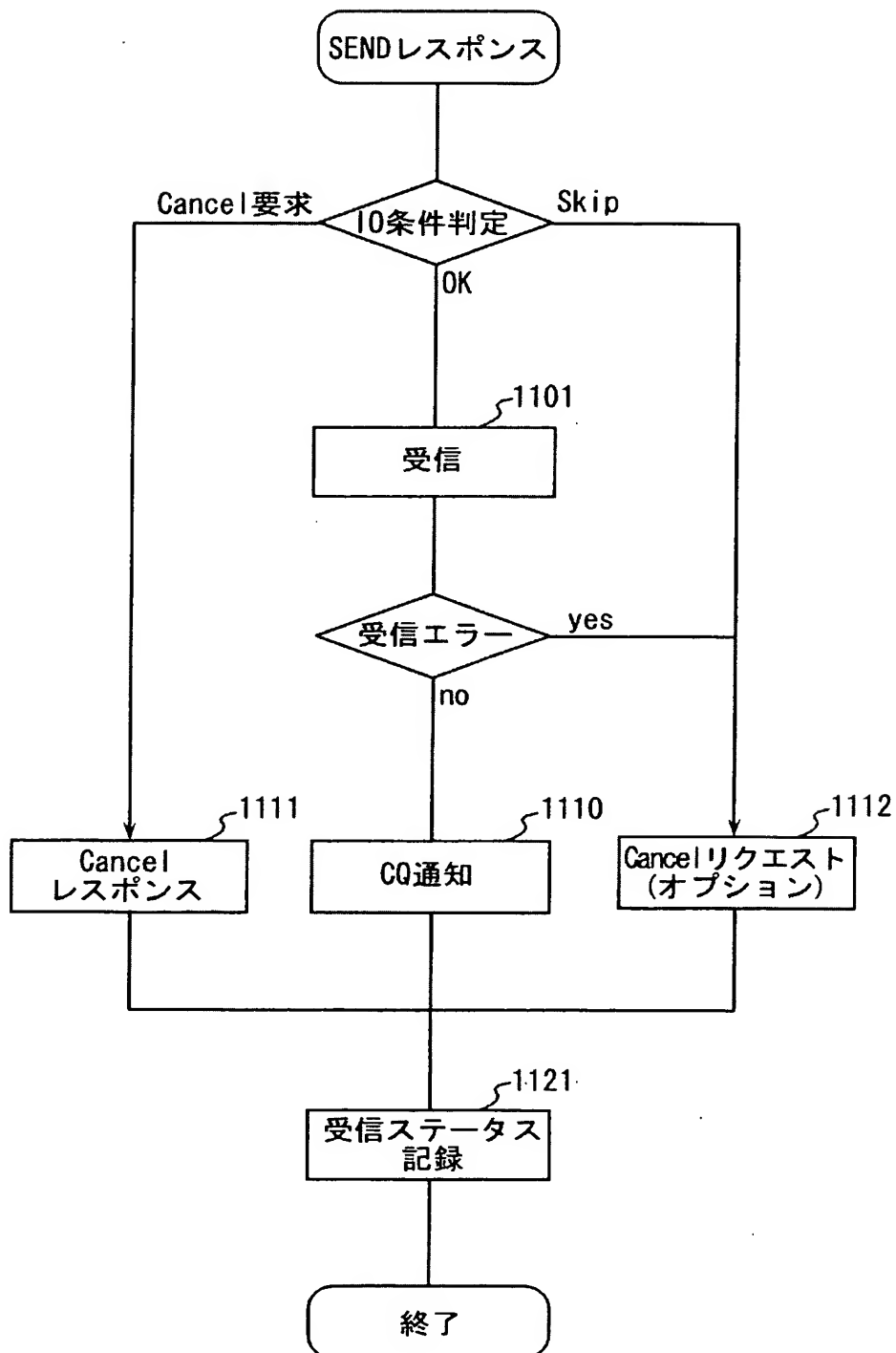
【図 11】

図 11



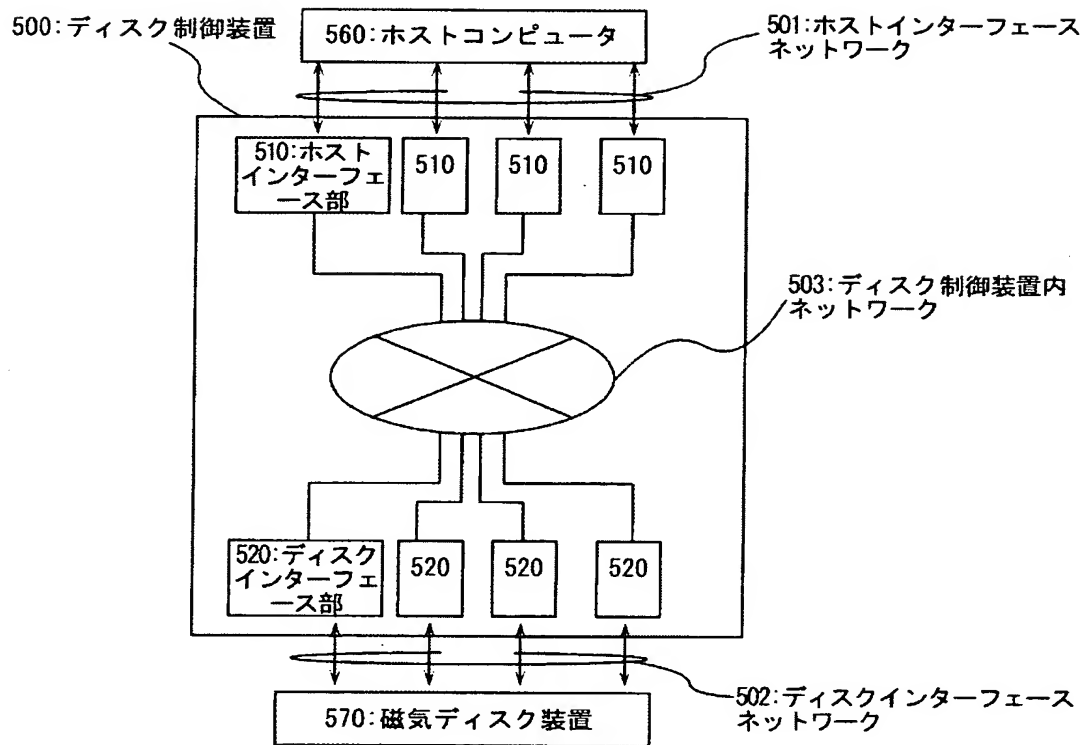
【図 12】

図 12



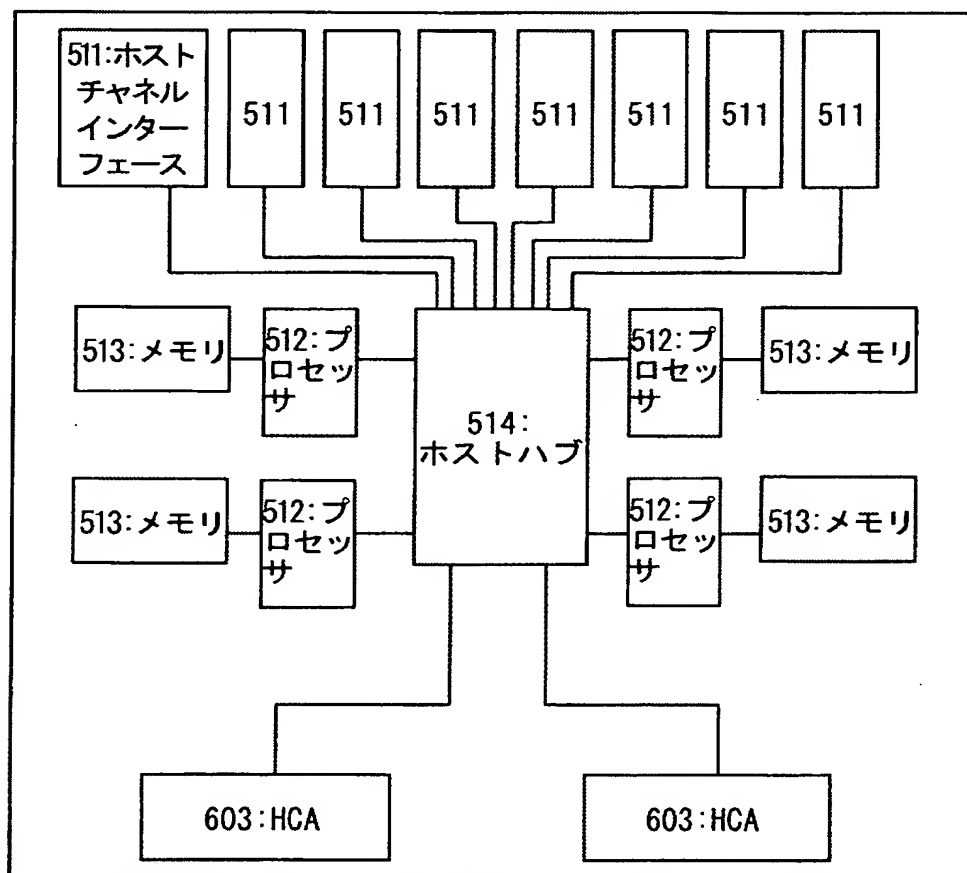
【図 13】

図 13



【図 14】

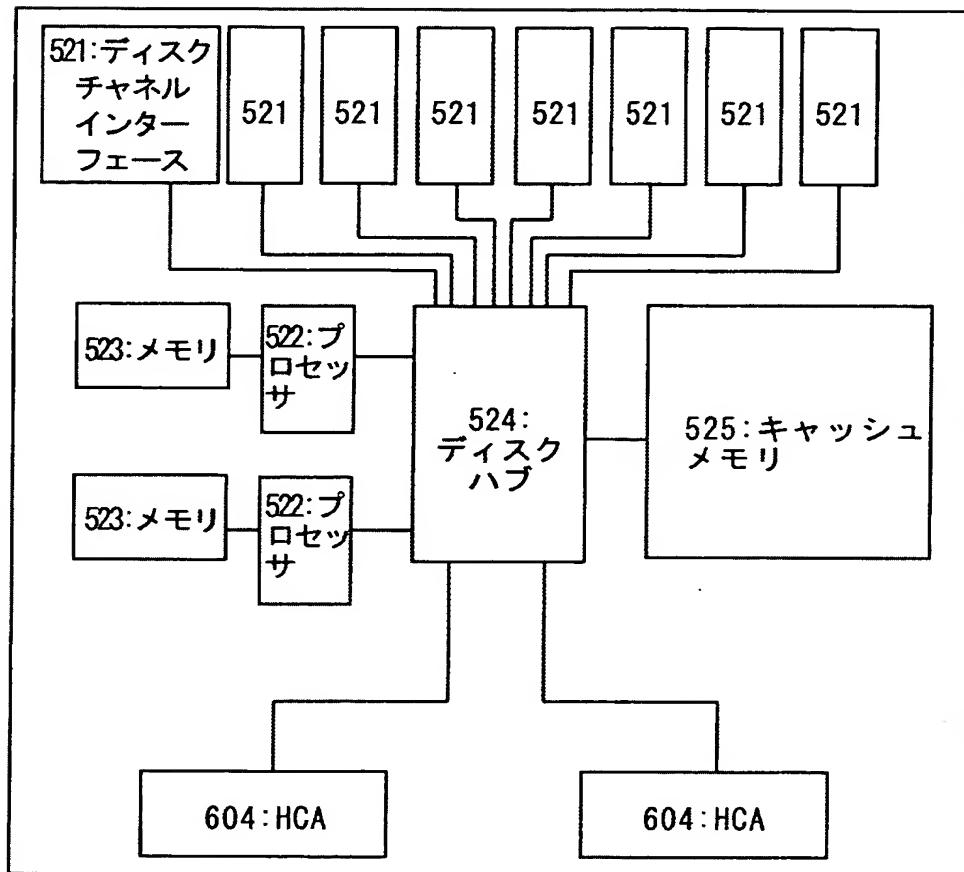
図 14





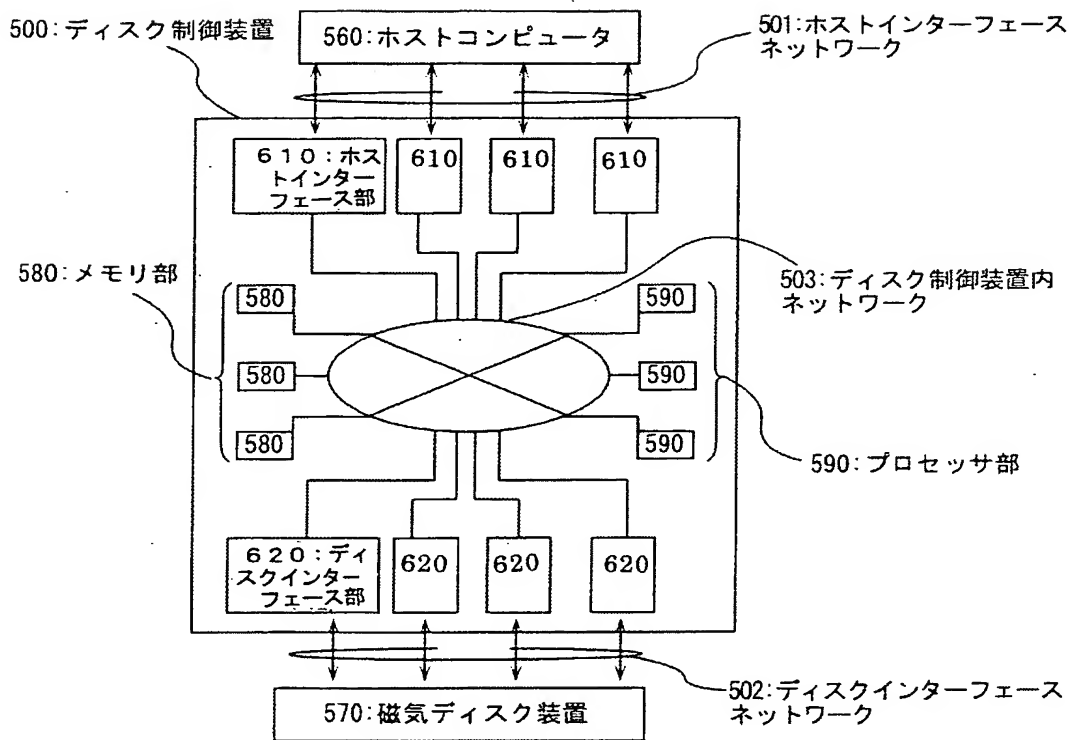
【図 15】

図 15



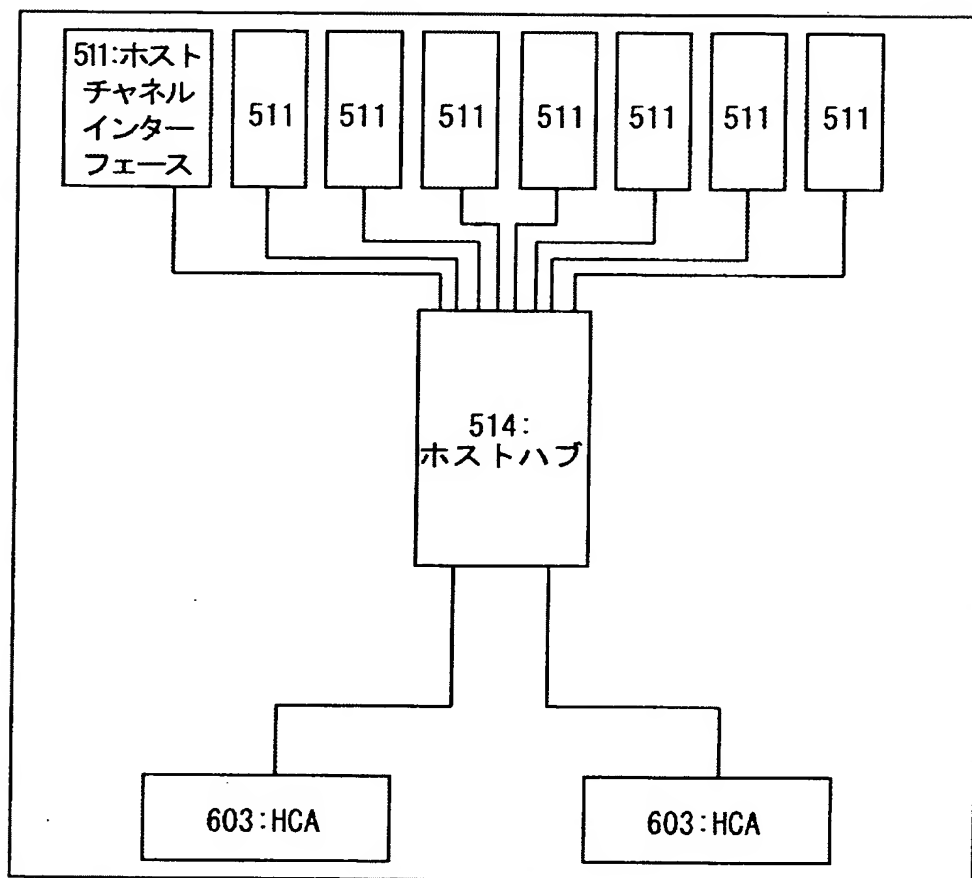
【図 16】

図 16



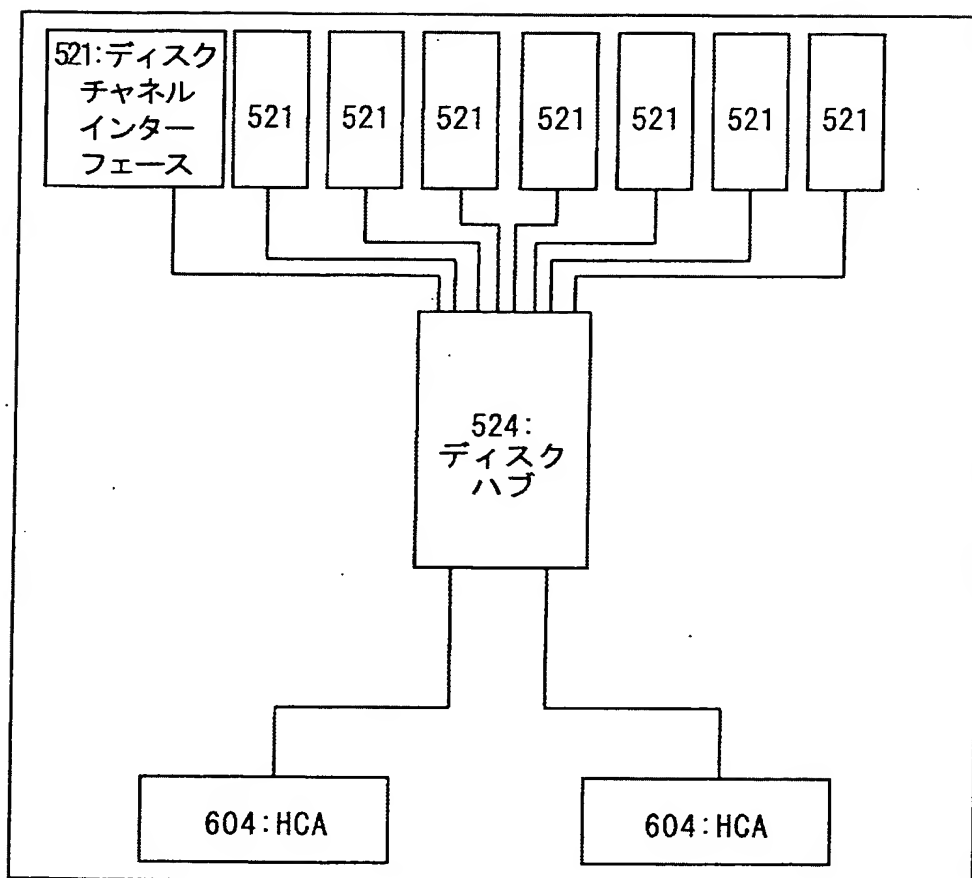
【図 17】

図 17



【図 18】

図 18



**【書類名】 要約書****【要約】**

**【課題】** 高い転送効率と高いアプリケーション処理効率を同時に実現するリライアブルデータ転送方法、および、これを用いたディスク制御装置を提供する。

**【解決手段】** イニシエータからターゲットにデータ転送を行う際、前記ターゲットで受信されたデータの妥当性を前記データに付加されているエラーチェックコードを用いて確認した上、前記妥当性を前記ターゲットから前記イニシエータに転送ステータスとして返送し、前記転送ステータスにより前記データ転送の際に転送エラーの発生していることが判明した場合、前記イニシエータから前記ターゲットに前記データの再送を行うリライアブル転送において、

前記イニシエータから前記ターゲットへの論理レコードのデータ転送方法であって、

前記イニシエータが発行した転送リクエストにより前記論理レコードの転送が正常に前記ターゲットに到着した時点で、前記論理レコードの前記転送リクエストに対応した完了ステータスを前記ターゲット内に存在する完了キューに通知し、

前記論理レコードを複数まとめて一括転送を行い、

前記イニシエータでは、前記一括転送の単位で前記転送ステータスの確認を行い、

前記ターゲットでは、前記論理レコードの前記転送リクエストに対応した完了ステータスを予め定められた一括転送条件に合致した前記論理レコードについて、その正常受信が完了した時点で前記ターゲット内に存在する完了キューに通知することを特徴とする。

**【選択図】** 図 1

特願 2 0 0 3 - 3 5 3 2 1 9

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 5 1 0 8 ]

1. 変更年月日  
[変更理由]

1 9 9 0 年 8 月 3 1 日  
新規登録

住 所  
氏 名

東京都千代田区神田駿河台 4 丁目 6 番地  
株式会社日立製作所